Coursework	**Spring 2019**
Statistical Pattern Recognition (COMP136)	**Roni Khardon**
Department of Computer Science, Tufts University	**Rishit Sheth**

Written Assignment 2

This assignment is due by the **start** of class on Tuesday, February 19.

1. (20 pts.) Consider an i.i.d. sample X of univariate random variables drawn from a Gaussian distribution with mean μ and precision λ (precision is the inverse of variance):

$$p(X|\mu, \lambda) = \prod_{i=1}^{N} \left(\frac{\lambda}{2\pi}\right)^{1/2} e^{-\frac{\lambda}{2}(x_i-\mu)^2}.$$

As described towards the end of Section 2.3.6 of the textbook (around equations 2.152-2.154), finding the *joint* posterior for the mean and the precision $p(\mu, \lambda|X)$ via conjugacy requires the Gaussian-Gamma distribution as prior:

$$p(\mu, \lambda|\mu_0, \beta_0, a_0, b_0) = \mathcal{N}(\mu|\mu_0, (\beta_0\lambda)^{-1})\text{Gamma}(\lambda|a_0, b_0).$$

Calculate the parameters $(\mu_N, \beta_N, a_N, b_N)$ of the posterior:

$$p(\mu, \lambda|X, \mu_N, \beta_N, a_N, b_N) = \mathcal{N}(\mu|\mu_N, (\beta_N\lambda)^{-1})\text{Gamma}(\lambda|a_N, b_N),$$

thereby verifying the conjugacy between likelihood and prior in this case. (Since the notation in Section 2.3.6 can be confusing, please use the notation here.)

2. (10 pts.) For the matrix

$$S = \begin{pmatrix} 5 & 0 & -4 \\ 0 & 4 & 0 \\ -4 & 0 & 5 \end{pmatrix},$$

(a) Develop the eigenvalue/eigenvector decomposition $V\Lambda V^T$ where V is the orthonormal matrix composed of eigenvectors, and

(b) Show how $z = (1, 2, 3)^\top$ can be expressed as a linear combination of eigenvectors, i.e., find scalars c_1, c_2, c_3 such that $z = \sum_{i=1}^{3} c_i v_i$ where $V = (v_1, v_2, v_3)$.

3. (10 pts.) Solve problem 3.7 (page 175) in the textbook.

4. (10 pts.) Solve problem 3.11 (page 175) in the textbook.

5. (10 pts.) Suppose we wish to infer the value of some unknown $x \in \mathbb{R}^2$ via a noisy observation. Specifically, let $p(x) = \mathcal{N}(x|\mu, \Sigma)$ where $\mu = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, and $\Sigma = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$. The noisy observation y is described by $p(y|x) = \mathcal{N}(y|2x_1 + x_2, 2)$. What is the probability of x given $y = 1$? How does the uncertainty in x change after y is observed? (Think about the dependency between the components of x and the "size" of the uncertainty.)